Refringence, field theory, and normal modes

Carlos Barceló*
Relativity and Cosmology group, School of Computer Science and Mathematics, Portsmouth University, Portsmouth PO1 2EG, Britain.

Stefano Liberati†
Physics Department, University of Maryland, College Park, MD 20742–4111, USA

Matt Visser‡
Physics Department, Washington University, Saint Louis MO 63130–4899, USA
(Dated: 19 November 2001; LaTeX-ed February 7, 2008)

In a previous paper [gr-qc/0104001; Class. Quant. Grav. **18** (2001) 3595-3610] we have shown that the occurrence of curved spacetime "effective Lorentzian geometries" is a generic result of linearizing an arbitrary classical field theory around some non-trivial background configuration. This observation explains the ubiquitous nature of the "analog models" for general relativity that have recently been developed based on condensed matter physics. In the simple (single scalar field) situation analyzed in our previous paper, there is a single unique effective metric; more complicated situations can lead to bi-metric and multi-metric theories. In the present paper we will investigate the conditions required to keep the situation under control and compatible with experiment — either by enforcing a unique effective metric (as would be required to be strictly compatible with the Einstein Equivalence Principle), or at the worst by arranging things so that there are multiple metrics that are all "close" to each other (in order to be compatible with the Eötvös experiment). The algebraically most general situation leads to a physical model whose mathematical description requires an extension of the usual notion of Finsler geometry to a Lorentzian-signature pseudo-Finsler geometry; while this is possibly of some interest in its own right, this particular case does not seem to be immediately relevant for either particle physics or gravitation. The key result is that wide classes of theories lend themselves to an effective metric description. This observation provides further evidence that the notion of "analog gravity" is rather generic.

PACS numbers: 04.40.-b; 04.60.-m; 11.10.-z; 45.20.-d; gr-qc/0111059
Keywords: Birefringence, analog gravity, field theory, normal modes

I. INTRODUCTION

Whenever you have a single classical scalar field governed by Lagrangian dynamics, the linearization of the field around any classical solution automatically provides you with a curved space "effective metric". Furthermore this effective metric has Lorentzian signature if and only if the PDE governing the scalar field is hyperbolic [1]. This observation, combined with the fact that quantization of the linearized fluctuations automatically provides an Einstein–Hilbert term in the one-loop effective action, led us to suggest that Einstein gravity is a generic low-energy limit of a wide class of quantum field theories [1].

A technical step in [1] that kept the calculation under control was to assume that one was dealing with a single scalar field, which automatically implied that there was a single unique effective metric. However, this assumption is very restrictive. By starting with a single scalar field it is not possible to reproduce the entire algebraic structure of the set of possible metrics in GR. Therefore, in the present article we turn our attention towards the somewhat messier general situation of multiple fields (multiple scalar fields, or a multi-component vector or tensor field). The key new results in this situation are:

1. In some situations the existence of a single unique effective metric can be deduced. This is the case that is compatible with strict application of the Einstein Equivalence Principle.

---

*Electronic address: carlos@hbar.wustl.edu; This work has been done at Washington University in Saint Louis supported by the Spanish MCYT; Currently supported by a European Community Marie Curie grant.
†Electronic address: liberati@physics.umd.edu; URL: `http://www2.physics.umd.edu/~liberati`; Supported by the US NSF
‡Electronic address: visser@kiwi.wustl.edu; URL: `http://www.physics.wustl.edu/~visser`; Supported by the US DOE

2. More generally there may be several distinct effective metrics on the same spacetime (this will generically occur if the dynamical system exhibits "birefringence" — more properly called "multi-refringence" — a situation in which different types of oscillation, typically referred to as different "polarizations", can propagate at different speed). As long as these various metrics are sufficiently "close" to each other, different types of matter will feel approximately the same geometry — as is required experimentally to satisfy the constraints deduced from the Eötvös experiment.

3. In the algebraically most general situation a mathematical structure is encountered which can best be viewed as a hyperbolic extension of the notion of a Finsler geometry; these pseudo-Finsler geometries bear the same relation to ordinary Finsler geometries that pseudo-Riemannian geometries (*aka* Lorentzian geometries) bear to Riemannian geometries—and many of the same sort of technical problems arise due to indefiniteness of the pseudo-Finslerian "metric". While these structures may be of interest in their own right, they do not seem to be immediately appropriate for either particle physics or gravitation.

In our previous paper we studied the single-scalar case in detail; that case being a perfect exemplar for the occurrence of a Lorentzian effective metric [1]. That paper also contained an extensive bibliography regarding analog models and we will be more selective in this present paper. We shall consider multi-component systems and push the "effective metric" analysis as far as practical, making extensive use of the theory of characteristic surfaces. We derive algebraic conditions that should be satisfied to keep the effective metric unique (or nearly so). We investigate the notion of "polarization" and the associated "Fresnel equation", leading to the notion of "multi-refringence". Finally we have a few words to say about the algebraically general pseudo-Finsler geometries — this is a subject that mathematically is very poorly developed, and we hope by this paper to generate some interest in what otherwise seems a rather arcane and abstract subject. The key message to take from the present effort is that while dealing with multiple fields is algebraically more tricky, there are nevertheless wide classes of dynamical systems that lend themselves to an effective metric description. Furthermore with multiple background fields, the background geometry is more flexible — the metric could in principle be completely general.

II. LINEARIZED FIELDS FROM ARBITRARY BACKGROUND SYSTEMS

A. Lagrangian analysis

Suppose we consider a *collection* of fields $\phi^A = \{\phi^1, \phi^2, \ldots\}$ whose dynamics is governed by some first-order Lagrangian $\mathcal{L}(\partial_\mu \phi^A, \phi^A)$. Here "first-order" is taken to mean that the Lagrangian depends only on the fields and their first derivatives. We want to consider linearized fluctuations around some background solution of the equations of motion. As in the single-field case [1] we write

$$\phi^A(t,\vec{x}) = \phi^A_0(t,\vec{x}) + \epsilon\phi^A_1(t,\vec{x}) + \frac{\epsilon^2}{2}\phi^A_2(t,\vec{x}) + O(\epsilon^3). \tag{1}$$

Now use this to expand the Lagrangian

$$\begin{aligned}\mathcal{L}(\partial_\mu \phi^A, \phi^A) &= \mathcal{L}(\partial_\mu \phi^A_0, \phi^A_0) + \epsilon\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\, \partial_\mu \phi^A_1 + \frac{\partial \mathcal{L}}{\partial \phi^A}\, \phi^A_1\right] + \frac{\epsilon^2}{2}\left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^A)}\, \partial_\mu \phi^A_2 + \frac{\partial \mathcal{L}}{\partial \phi^A}\, \phi^A_2\right] \\ &+ \frac{\epsilon^2}{2}\left[\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\,\partial(\partial_\nu \phi^B)}\, \partial_\mu \phi^A_1\, \partial_\nu \phi^B_1 + 2\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\,\partial \phi^B}\, \partial_\mu \phi^A_1\, \phi^B_1 + \frac{\partial^2 \mathcal{L}}{\partial \phi^A\,\partial \phi^B}\, \phi^A_1\, \phi^B_1\right] + O(\epsilon^3).\end{aligned} \tag{2}$$

Consider the action

$$S[\phi^A] = \int \mathrm{d}^{d+1}x\; \mathcal{L}(\partial_\mu \phi^A, \phi^A). \tag{3}$$

Doing so allows us to integrate by parts. As in the single-field case [1] we can use the Euler–Lagrange equations to discard the linear terms (since we are linearizing around a solution of the equations of motion) and so get

$$\begin{aligned}S[\phi^A] &= S[\phi^A_0] \\ &+ \frac{\epsilon^2}{2}\int \mathrm{d}^{d+1}x\left[\left\{\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\,\partial(\partial_\nu \phi^B)}\right\} \partial_\mu \phi^A_1\, \partial_\nu \phi^B_1 + 2\left\{\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\,\partial \phi^B}\right\} \partial_\mu \phi^A_1\, \phi^B_1 + \left\{\frac{\partial^2 \mathcal{L}}{\partial \phi^A\,\partial \phi^B}\right\} \phi^A_1\, \phi^B_1\right] \\ &+ O(\epsilon^3).\end{aligned} \tag{4}$$

Because the fields now carry indices (AB) we cannot cast the action into quite as simple a form as was possible in the single-field case. The equation of motion for the linearized fluctuations are now read off as

$$\partial_\mu \left(\left\{ \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial(\partial_\nu \phi^B)} \right\} \partial_\nu \phi_1^B \right) + \partial_\mu \left(\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial \phi^B}\; \phi_1^B \right) - \partial_\mu \phi_1^B\; \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^B)\, \partial \phi^A} - \left(\frac{\partial^2 \mathcal{L}}{\partial \phi^A\, \partial \phi^B} \right) \phi_1^B = 0. \quad (5)$$

This is a linear second-order *system* of partial differential equations with position-dependent coefficients. This system of PDEs is automatically self-adjoint (with respect to the trivial "flat" measure $\mathrm{d}^{d+1}x$).

To simplify the notation we introduce a number of definitions. First

$$f^{\mu\nu}{}_{AB} \equiv \frac{1}{2} \left(\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial(\partial_\nu \phi^B)} + \frac{\partial^2 \mathcal{L}}{\partial(\partial_\nu \phi^A)\, \partial(\partial_\mu \phi^B)} \right). \quad (6)$$

This quantity is independently symmetric under interchange of μ, ν and A, B. In the case of Bose fields we will want to interpret this as something related to some sort of "metric", but the interpretation is not as straightforward as for the single-field case.

Next, define

$$\Gamma^\mu{}_{AB} \equiv + \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial \phi^B} - \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^B)\, \partial \phi^A} + \frac{1}{2} \partial_\nu \left(\frac{\partial^2 \mathcal{L}}{\partial(\partial_\nu \phi^A)\, \partial(\partial_\mu \phi^B)} - \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial(\partial_\nu \phi^B)} \right). \quad (7)$$

This quantity is anti-symmetric in A, B. For Bose fields we will want to interpret this as some sort of "connexion". Equivalently we could write

$$\Gamma^\mu{}_{AB} \equiv + \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial \phi^B} - \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^B)\, \partial \phi^A} - \frac{1}{2} \partial_\nu \left(\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial(\partial_\nu \phi^B)} - \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^B)\, \partial(\partial_\nu \phi^A)} \right). \quad (8)$$

It is useful to note that

$$\partial_\mu\, \Gamma^\mu{}_{AB} = \partial_\mu\, \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial \phi^B} - \partial_\mu\, \frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^B)\, \partial \phi^A}. \quad (9)$$

In the case of Fermi fields one typically has $f^{\mu\nu}{}_{AB} = 0$, while the $\Gamma^\mu{}_{AB}$ are most usefully thought of as generalizations of the Dirac matrices. Finally, define

$$K_{AB} = -\frac{\partial^2 \mathcal{L}}{\partial \phi^A\, \partial \phi^B} + \frac{1}{2} \partial_\mu \left(\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^A)\, \partial \phi^B} \right) + \frac{1}{2} \partial_\mu \left(\frac{\partial^2 \mathcal{L}}{\partial(\partial_\mu \phi^B)\, \partial \phi^A} \right). \quad (10)$$

This quantity is by construction symmetric in (AB). We will want to interpret this as some sort of "potential" or "mass matrix". Then the crucial point for the following discussion is to realize that equation (5) can be written in the form

$$\partial_\mu \left(f^{\mu\nu}{}_{AB}\; \partial_\nu \phi_1^B \right) + \frac{1}{2} \left[\Gamma^\mu_{AB}\; \partial_\mu \phi_1^B + \partial_\mu (\Gamma^\mu_{AB}\; \phi_1^B) \right] + K_{AB}\; \phi_1^B = 0. \quad (11)$$

Now it is more transparent that this is a formally self-adjoint second-order linear *system* of PDEs.

B. Systems of second-order hyperbolic PDEs

One can arrive at a similar set of equations by directly linearizing an arbitrary system of second-order PDEs, that is, without relying on the existence of a Lagrangian dynamics. Start by considering an arbitrary system of second-order PDEs written in the form

$$G_A(x, \phi^B, \partial_\mu \phi^B, \partial_\mu \partial_\nu \phi^B) = 0. \quad (12)$$

The PDE does not have to be linear or even quasi-linear; there are as many equations as there are fields. As in the single-field case defining hyperbolicity for such a general equation is not easy — not even Courant and Hilbert [2] consider this case explicitly. Indeed we shall adapt their discussion and shall define hyperbolicity in terms of the linearized equation: Suppose we linearize around some solution ϕ_0^A, writing

$$\phi^A(t, \vec{x}) = \phi_0^A(t, \vec{x}) + \epsilon \phi_1^A(t, \vec{x}) + O(\epsilon^2). \quad (13)$$

Then

$$\frac{\partial G_A}{\partial(\partial_\mu \partial_\nu \phi^B)}\; \partial_\mu \partial_\nu \phi_1^B + \frac{\partial G_A}{\partial(\partial_\mu \phi^B)}\; \partial_\mu \phi_1^B + \frac{\partial G_A}{\partial \phi^B}\; \phi_1^B = 0. \tag{14}$$

That is: the fluctuation satisfies a second-order system of linear PDEs with time-dependent and position-dependent coefficients (these coefficients depend on the background field you are linearizing around). This equation can be written in a manner similar to (11)

$$\partial_\mu \left(\tilde{f}^{\mu\nu}{}_{AB}\; \partial_\nu \phi_1^B \right) + \tilde{\Gamma}^\mu_{AB}\; \partial_\mu \phi_1^B + \tilde{K}_{AB}\; \phi_1^B = 0, \tag{15}$$

by defining

$$\tilde{f}^{\mu\nu}{}_{AB} = \frac{\partial G_A}{\partial(\partial_\mu \partial_\nu \phi^B)}, \tag{16}$$

$$\tilde{\Gamma}^\mu_{AB} = \frac{\partial G_A}{\partial(\partial_\mu \phi^B)} - \partial_\nu \left[\frac{\partial F_A}{\partial(\partial_\mu \partial_\nu \phi^B)} \right], \tag{17}$$

$$\tilde{K}_{AB} = \frac{\partial G_A}{\partial \phi^B}. \tag{18}$$

Though the form of the coefficients is somewhat different from that appearing in the Lagrangian-based analysis (in general the equation is not automatically self-adjoint), much of the discussion that follows can be applied (with minor modifications) to both cases.

C. Eikonal approximation

The theory of *systems* of second-order PDE is relatively complicated and much less transparent than that for a *single* second-order PDE. (See, for example, Courant and Hilbert [2], volume 2, pp 577–618, or the Encyclopedic Dictionary of Mathematics [3].) We would like to be able to construct some notion of spacetime metric directly from the quantities $f^{\mu\nu}{}_{AB}$ (or $\tilde{f}^{\mu\nu}{}_{AB}$), but we shall soon see that doing so is a somewhat tricky proposition.

Consider an eikonal approximation for an arbitrary direction in field space, that is, take

$$\phi^A(x) = \epsilon^A(x)\; \exp[-i\varphi(x)], \tag{19}$$

with $\epsilon^A(x)$ a slowly varying amplitude, and $\varphi(x)$ a rapidly varying phase. In this eikonal approximation (where we neglect gradients in the amplitude, and gradients in the coefficients of the PDEs, retaining only the gradients of the phase) the linearized system of PDEs (11) becomes

$$\{ f^{\mu\nu}{}_{AB}\; \partial_\mu \varphi(x)\; \partial_\nu \varphi(x) + \Gamma^\mu{}_{AB}\; \partial_\mu \varphi(x) + K_{AB} \}\; \epsilon_1^B = 0. \tag{20}$$

This has a nontrivial solution if and only if $\epsilon^A(x)$ is a null eigenvector of the matrix

$$f^{\mu\nu}{}_{AB}\; \partial_\mu \varphi(x)\; \partial_\nu \varphi(x) + \Gamma^\mu{}_{AB}\; \partial_\mu \varphi(x) + K_{AB}. \tag{21}$$

Now, the condition for such a null eigenvector to exist is that

$$F(p,q) \equiv \det \{ f^{\mu\nu}{}_{AB}\; \partial_\mu \varphi(x)\; \partial_\nu \varphi(x) + \Gamma^\mu{}_{AB}\; \partial_\mu \varphi(x) + K_{AB} \} = 0, \tag{22}$$

with the determinant to be taken on the field space indices. This is the natural generalization to the current situation of the Fresnel equation of bi-refringent optics [4, 5]. Following the analogy with the situation in electrodynamics (either nonlinear electrodynamics, or more prosaically propagation in a bi-refringent crystal), the null eigenvector $\epsilon^A(x)$ would correspond to a specific "polarization". The Fresnel equation then describes how different polarizations can propagate at different velocities (or in the language to be used later in the paper, can see different metric structures). In particle physics language this determinant condition $F(p,q) = 0$ is the natural generalization of the "mass shell" constraint. Indeed it is useful to define the mass shell as a subset of the cotangent space by

$$\mathcal{F}(q) \equiv \left\{ p_\mu \;\middle|\; F(p,q) = 0 \right\}. \tag{23}$$

In more mathematical language we are looking at the null space of the determinant of the "symbol" of the system of PDEs. By investigating $F(p,q)$ one can recover part (not all) of the information encoded in the matrices $f^{\mu\nu}{}_{AB}$, $\Gamma^{\mu}{}_{AB}$, and K_{AB}, or equivalently in the "generalized Fresnel equation" (22). (Note that for the determinant equation to be useful it should be non-vacuous; in particular one should carefully eliminate all gauge and spurious degrees of freedom before constructing this "generalized Fresnel equation", since otherwise the determinant will be identically zero. Some examples of this phenomenon will be given later in the paper, see subsection IV E). We now want to make this analogy with optics more precise, by carefully considering the notion of characteristics and characteristic surfaces. We will see how to extract from the the high-frequency high-momentum regime described by the eikonal approximation all the information concerning the causal structure of the theory.

III. CAUSAL STRUCTURES

One of the key structures that a Lorentzian spacetime metric provides is the notion of causal relationships. This suggests that it may be profitable to try to work backwards from the causal structure to determine a Lorentzian metric. Now the causal structure implicit in the system of second-order PDEs given in equation (11) is described in terms of the characteristic surfaces, and it is for this reason that we now focus on characteristics as a way of encoding causal structure, and as a surrogate for some notion of Lorentzian metric. Note that via the Hadamard theory of surfaces of discontinuity the characteristics can be identified with the infinite-momentum limit of the eikonal approximation [6]. That is, when extracting the characteristic surfaces we neglect subdominant terms in the generalized Fresnel equation and focus only on the leading term in the symbol ($f^{\mu\nu}{}_{AB}$). In particle physics language going to the infinite-momentum limit puts us on the light cone instead of the mass shell; and it is the light cone that is more useful in determining causal structure. The "normal cone" at some specified point q, consisting of the locus of normals to the characteristic surfaces, is defined by

$$\mathcal{N}(q) \equiv \left\{ p_\mu \;\middle|\; \det\left(f^{\mu\nu}{}_{AB}\;\; p_\mu\, p_\mu\right) = 0 \right\}. \tag{24}$$

As was the case for the Fresnel equation (22), the determinant is to be taken on the field indices AB. Remember to eliminate spurious and gauge degrees of freedom so that this determinant is not identically zero. (See Courant and Hilbert [2], volume 2 "Partial differential equations", page 580.) We emphasise that the algebraic equation defining the normal cone is the leading term in the Fresnel equation encountered in discussing the eikonal approximation. If there are N fields in total then this "normal cone" will generically consist of N nested sheets each with the topology (not necessarily the geometry) of a cone. Often several of these cones will coincide, which is not particularly troublesome, but unfortunately it is also common for some of these cones to be degenerate, which is more problematic. See Courant and Hilbert footnote 1 on page 592:

> *It may be remarked that the present state of the theory of algebraic surfaces does not permit entirely satisfactory applications to the questions of reality of geometric structures which confront us here.*

Note that if one is dealing with Fermi fields where $f^{\mu\nu}{}_{AB} = 0$, one has a choice: either iterate the first-order system of PDEs to produce a second-order system and then apply the previous logic, or go to the highest nontrivial term in the symbol and rephrase the discussion below (*mutatis mutandis*) in terms of

$$\mathcal{N}(q) \equiv \left\{ p_\mu \;\middle|\; \det\left(\Gamma^{\mu}{}_{AB}\;\; p_\mu\right) = 0 \right\}. \tag{25}$$

Returning to the Bosonic case, it is convenient to define a function $Q(q,p)$ on the co-tangent bundle

$$Q(q,p) \equiv \det\left(f^{\mu\nu}{}_{AB}(q)\; p_\mu\, p_\mu\right). \tag{26}$$

The function $Q(q,p)$ defines a completely-symmetric spacetime tensor (actually, a tensor density) with $2N$ indices

$$Q(q,p) = Q^{\mu_1\nu_1\mu_2\nu_2\cdots\mu_N\nu_N}(q)\;\; p_{\mu_1}\, p_{\nu_1}\, p_{\mu_2}\, p_{\nu_2} \cdots\, p_{\mu_N}\, p_{\nu_N}. \tag{27}$$

(Remember that $f^{\mu\nu}{}_{AB}$ is symmetric in both $\mu\nu$ and AB independently.) Explicitly, using the expansion of the determinant in terms of completely antisymmetric field-space Levi–Civita tensors

$$\begin{aligned} Q^{\mu_1\nu_1\mu_2\nu_2\cdots\mu_N\nu_N}(q) \;=\; & \frac{1}{N!}\, \epsilon^{A_1A_2A_3\cdots A_N}\, \epsilon^{B_1B_2B_3\cdots B_N} \\ & f^{\mu_1\nu_1}{}_{A_1B_1}\, f^{\mu_2\nu_2}{}_{A_2B_2} \cdots f^{\mu_N\nu_N}{}_{A_NB_N}. \end{aligned} \tag{28}$$

In terms of this $Q(q,p)$ function the normal cone is

$$\mathcal{N}(q) \equiv \left\{ p_\mu \;\middle|\; Q(q,p) = 0 \right\}. \tag{29}$$

In contrast, the "Monge cone" (*aka* "ray cone", *aka* "characteristic cone", *aka* "null cone") is the envelope of the set of characteristic surfaces through the point q. Thus the "Monge cone" is dual to the "normal cone", its explicit construction is given by (Courant and Hilbert [2], volume 2, pp 583):

$$\mathcal{M}(q) = \left\{ t^\mu = \frac{\partial Q(q,p)}{\partial p_\mu} \;\middle|\; p_\mu \in \mathcal{N}(q) \right\}. \tag{30}$$

Equivalently

$$\mathcal{M}(q) = \left\{ t^\mu = Q^{\mu\mu_2\cdots\mu_{2N}}(q)\, p_{\mu_2}\, p_{\mu_3} \cdots\, p_{\mu_{2N}} \;\middle|\; p_\mu \in \mathcal{N}(q) \right\}. \tag{31}$$

Unfortunately, even if the normal cone is pleasantly behaved, the Monge cone (null cone) may be quite messy. Indeed Courant and Hilbert remark:

> *Even if [the normal cone] is a relatively simple algebraic cone of degree [2N], the ray cone [Monge cone/null cone] may have singularities, or isolated rays, and need not consist of separate smooth conical shells.*

(Courant and Hilbert [2], volume 2, p 584). Defining the notion of "hyperbolicity" in this general context boils down to the question of just what constraints have to be placed on the original system of PDEs to make sure the Monge cone is well-behaved.

For completeness, we mention a general construction for finding bi-characteristic rays (these are lines which are guaranteed to lie on characteristic surfaces). Start by picking some point q and a co-tangent vector p which is in the normal cone $\mathcal{N}(q)$. Now with these initial conditions solve the "Hamiltonian equation"

$$\frac{dq^\mu}{ds} = -\frac{\partial Q(q,p)}{\partial p_\mu}; \qquad \frac{dp_\mu}{ds} = +\frac{\partial Q(q,p)}{\partial q^\mu}. \tag{32}$$

Here s is just a parameter, it is not physical time. The resulting curve $[q^\mu(s),\, p_\mu(s)]$ is called a bi-characteristic ray. For all s we have $p_\mu(s) \in \mathcal{N}(q(s))$ and $dq^\mu/ds \in \mathcal{M}(q(s))$. (The momentum lies on the normal cone, and the velocity lies on the Monge cone [null cone].)

In a manner similar to the way in which the Monge cone is the dual of the normal cone, we can define a dual to the mass shell, this dual now being a subset of the tangent space

$$\mathcal{F}^*(q) \equiv \left\{ \frac{\partial F(p,q)}{\partial p_\mu} \;\middle|\; p \in \mathcal{F}(q) \right\}. \tag{33}$$

In simple situations where there is a single metric the Monge cone can be constructed simply by "raising the index" in the definition of the normal cone, and similarly for the dual mass shell in terms of the mass shell.

The structure of the normal and Monge cones encode all the information related with the causal propagation of signals associated with the system of PDEs. In next section, we will detail how to relate this causal structure with the existence of some spacetime metrics in different specific situations, from the experimentally favored single-metric theory deduced from the Equivalence Principle to the most complicated case of pseudo-Finsler geometries.

IV. GEOMETRICAL INTERPRETATION

A. Field redefinitions

We have seen that the causal structure of a system of coupled PDEs of the form (11) could in general be rather complicated. However, there are particular cases in which it is relatively easy to find a geometrical interpretation of what is happening. First, it is important to realize that we are always free to perform a field redefinition

$$\phi^A \to \bar{\phi}^A = h^A(\phi^B), \tag{34}$$

essentially a coordinate change in field space. This induces a linear transformation on the linearized fields ϕ_1^A,

$$\phi_1^A \to \bar{\phi}_1^A = \left.\frac{\partial h^A}{\partial \phi^B}\right|_{\phi_0^C} \phi_1^B = [L^{-1}(\phi_0^C)]^A{}_B \;\; \phi_1^B, \tag{35}$$

and, therefore, a redefinition of the $f^{\mu\nu}{}_{AB}$, $\Gamma^\mu{}_{CD}$, and K_{AB}. It is convenient to adopt matrix notation, suppressing the field indices (but keeping the spacetime indices explicit). In matrix notation

$$\boldsymbol{\phi}_1 \to \bar{\boldsymbol{\phi}}_1 = \boldsymbol{L^{-1}}\; \boldsymbol{\phi}_1, \tag{36}$$

while the self-adjoint system of PDEs (11) can be written in the form

$$\partial_\mu \left(\boldsymbol{f}^{\mu\nu}\; \partial_\nu \boldsymbol{\phi}_1\right) + \boldsymbol{\Gamma}^\mu\; \partial_\mu \boldsymbol{\phi}_1 + \frac{1}{2}\partial_\mu(\boldsymbol{\Gamma}^\mu)\; \boldsymbol{\phi}_1 + \boldsymbol{K}\; \boldsymbol{\phi}_1 = 0. \tag{37}$$

A brief computation yields

$$\boldsymbol{f}^{\mu\nu} \;\; \to \;\; \bar{\boldsymbol{f}}^{\mu\nu} = \boldsymbol{L}^T\; \boldsymbol{f}^{\mu\nu}\; \boldsymbol{L}; \tag{38}$$

$$\boldsymbol{\Gamma}^\mu \;\; \to \;\; \bar{\boldsymbol{\Gamma}}^\mu = \boldsymbol{L}^T\; \boldsymbol{\Gamma}^\mu\; \boldsymbol{L} + \boldsymbol{L}^T\; \boldsymbol{f}^{\mu\nu}\; \partial_\nu \boldsymbol{L} - \partial_\nu \boldsymbol{L}^T\; \boldsymbol{f}^{\mu\nu}\; \boldsymbol{L}; \tag{39}$$

$$\begin{aligned}\boldsymbol{K} \;\; \to \;\; \bar{\boldsymbol{K}} = \boldsymbol{L}^T\; \boldsymbol{K}\; \boldsymbol{L} - \frac{1}{2}\left(\partial_\nu \boldsymbol{L}^T\; \boldsymbol{\Gamma}^\nu\; \boldsymbol{L} - \boldsymbol{L}^T\; \boldsymbol{\Gamma}^\nu\; \partial_\nu \boldsymbol{L}\right) + \frac{1}{2}\partial_\mu \left(\boldsymbol{L}^T\; \boldsymbol{f}^{\mu\nu}\; \partial_\nu \boldsymbol{L} + \partial_\nu \boldsymbol{L}^T\; \boldsymbol{f}^{\mu\nu}\; \boldsymbol{L}\right) \\ -\partial_\nu \boldsymbol{L}^T\; \boldsymbol{f}^{\mu\nu}\; \partial_\mu \boldsymbol{L}.\end{aligned} \tag{40}$$

Note that because of the matrix multiplications the ordering is important. Also note that $(\boldsymbol{f}^{\mu\nu})^T = \boldsymbol{f}^{\mu\nu}$, while $(\boldsymbol{\Gamma}^\mu)^T = -\boldsymbol{\Gamma}^\mu$ and $(\boldsymbol{K})^T = \boldsymbol{K}$, where the transpose is on the field indices (not the spacetime indices), and that these symmetry properties are preserved under these background-field-dependent coordinate transformations. The point is that we can use these field redefinitions to simplify the matrix $f^{\mu\nu}{}_{AB}$ using position-dependent general linear transformations — this will be useful below. (For our purposes, the detailed form of the non-homogeneous derivative terms is often not important, though the fact of their presence is.)

B. Einstein Equivalence Principle

The physically simplest situation one can encounter is when the multiple fields being analyzed simply correspond to different but equivalent polarizations of a single field. By looking at the eikonal approximation of section II A, one can verify if there exist multiple independent polarizations that are solutions of equation (20) for the *same* phase function φ. Those multiple solutions should be grouped into a single class, all having the same propagation features. The number of independent solutions grouped into equivalent classes will give us information about the degree of degeneracy of the matrix $f^{\mu\nu}{}_{AB}$. (Here we are thinking of this object as a matrix in field space, whose elements are spacetime matrices).

The simplest possible situation regarding the geometrical structure emerging from a field theoretical mode analysis is that in which all linearized fields can be grouped into a single class. This implies that there must be some choice of field variables so that all the new $\bar{\phi}_1^A$ see the same metric, that is:

$$\bar{f}^{\mu\nu}{}_{AB} = \delta_{AB}\; f^{\mu\nu} = \delta_{AB}\; \sqrt{-g}\; g^{\mu\nu}. \tag{41}$$

This corresponds to the field theoretical analog PDE systems obeying strict adherence to the Einstein Equivalence Principle. If we use any other choice of field variables then we must have

$$f^{\mu\nu}{}_{AB} = h_{AB}\; f^{\mu\nu} = h_{AB}\; \sqrt{-g}\; g^{\mu\nu}. \tag{42}$$

This "factorization" condition on $f^{\mu\nu}{}_{AB}$ is a necessary and sufficient condition for strict adherence to the Einstein Equivalence Principle. Indeed for strict adherence to the Einstein Equivalence Principle you would also want to demand that in the $\bar{\phi}_1^A$ field variables $\bar{\Gamma}^\mu{}_{AB} = 0$, and one would want the "mass matrix" $\bar{K}_{AB}$ to be position independent. (At the very worst $\bar{K}_{AB}$ might contain curvature coupling terms that would go to zero in the flat space limit.)

In the usual formulation of general relativity, strict adherence to the Einstein Equivalence Principle is enforced by a policy of "minimal substitution" — the matter Lagrangian is taken to be a flat Minkowski-space Lagrangian with

the substitution $\eta_{\mu\nu} \to g_{\mu\nu}$. Since, by fiat, there is only one metric put into this Lagrangian, a unique spacetime metric will emerge for all the matter fields. (At least as long as all "external fields" are set to zero.) While the Einstein Equivalence Principle is certainly compatible with experiment, strictly enforcing it may be overkill — indeed the low-energy field theories arising from string theory and its relatives (in particular, various scalar tensor theories) do not strictly adhere to the Einstein Equivalence Principle, so it is useful to keep a little flexibility on this point.

C. Multiple metrics

Suppose now that we are looking for a multi-metric theory, then there must be some choice of field variables so that all the linearized fields $\bar{\phi}_1^A$ "decouple" and see independent metrics. That is, there must be a "diagonal" representation in field space such that

$$\bar{f}^{\mu\nu}{}_{AB} = \text{diag}\{f_1^{\mu\nu}, f_2^{\mu\nu}, f_3^{\mu\nu}, \cdots f_N^{\mu\nu}\} = \text{diag}\{\sqrt{-g_1}\; g_1^{\mu\nu},\; \sqrt{-g_2}\; g_2^{\mu\nu},\; \sqrt{-g_3}\; g_3^{\mu\nu}, \cdots \sqrt{-g_N}\; g_N^{\mu\nu}\}. \tag{43}$$

If we use any other choice of field variables then the necessary and sufficient condition for the $f^{\mu\nu}{}_{AB}$ to be simultaneously diagonalizable in field space is that $\forall\ \mu, \nu, \alpha, \beta$:

$$f^{\mu\nu}{}_{AB}\; f^{\alpha\beta}{}_{BC} = f^{\alpha\beta}{}_{AB}\; f^{\mu\nu}{}_{BC}; \qquad \text{that is} \qquad [\, \boldsymbol{f}^{\mu\nu}, \boldsymbol{f}^{\alpha\beta}\,] = 0. \tag{44}$$

Subject to this constraint, after diagonalizing $f^{\mu\nu}{}_{AB}$ in field space, each field couples to its own independent spacetime metric. If the original system of PDEs [equation (11)] is hyperbolic then all these metrics will have Lorentzian signature. Note that this is the behaviour that typically occurs in nonlinear electrodynamics and in birefringent crystals; and note that such birefringence is *not* in conflict with the physical intent of the Einstein Equivalence Principle because in both of these situations there is an external field (the electromagnetic field; the rest frame of the crystal itself) that softly breaks the Lorentz invariance. One thing that we can definitely say based on the Eötvös experiment [the observed universality of free fall] is that those low-energy quantum fields relevant to describing ordinary bulk matter all see approximately the same effective metric.

Note that the case of multiple metrics as defined above is related to (though not identical to) the notion of "reducibility" as introduced by Courant and Hilbert [2], p 596. The function $Q(q,p)$ is said to be "reducible" when it *factorizes* into lower-order polynomials. If it factorizes completely into simple quadratic products then:

$$Q(q,p) = \prod_{A=1}^{N} Q_A(q,p) = \prod_{A=1}^{N} \left(h_A^{\mu\nu}(q)\; p_\mu\; p_\mu\right). \tag{45}$$

If this "quadratic reducibility" property is satisfied then the normal cone consists of N nested topological cones each of which is geometrically a cone, and we can define a variant "multiple-metric theory" by defining the N spacetime metrics using the prescription

$$\sqrt{-g_A}\;\; g_A^{\mu\nu} = h_A^{\mu\nu}. \tag{46}$$

Even then, in order for the system (11) to be deemed to be hyperbolic, one must additionally insist that *each* of these matrices have Lorentzian signature and that the causal structures derived from each of these Lorentzian metric be compatible with each other: not only should each Lorentzian metric satisfy some sort of "chronology condition" or "causality condition" (no closed null or timelike paths), but to prevent an ill-posed problem one must insist that one cannot even form closed causal loops by using a chain of causal line segments belonging to different Lorentzian metrics.

Now if we have a multiple metric theory defined in terms of simultaneously diagonalizing the $f^{\mu\nu}{}_{AB}$ this automatically satisfies quadratic reducibility for $Q(q,p)$. The converse is not necessarily true, and the class of multiple-metric theories defined in terms of quadratic reducibility is more general than that defined in terms of simultaneous diagonalization of the kinetic terms. (This last comment is not supposed to be obvious; we shall give an explicit example of this phenomenon when we discuss the generic two-field situation in section V.)

In many cases all these metrics will coincide, and we can speak of *the* spacetime metric; this happens for instance (modulo some technical issues such as gauge fixing) for both the Maxwell equation and the Dirac equation. In terms of $Q(q,p)$ this requires

$$Q(q,p) = [Q_0(q,p)]^N = \left[\, h_0^{\mu\nu}(q)\; p_\mu\; p_\mu\, \right]^N. \tag{47}$$

D. The general case: pseudo-Finsler geometries

The most general case, when ${f^{\mu\nu}}_{AB}$ cannot be described in terms of a single metric, or even multiple metrics, must be dealt with by the formalism of pseudo-Finsler geometries. What we call pseudo-Finsler geometries bear the same relation to Finsler geometries that pseudo-Riemannian geometries (*aka* Lorentzian geometries) bear to Riemannian geometries—and they exhibit many of the same sort of technical problems that arise due to indefiniteness of the "metric".

Remember that for Riemannian geometries all the physically interesting quantities (metric, Riemann tensor, *etc.*) can be encoded in terms of coincidence limits of the (real) distance function $d(x, y)$ and its derivatives. If we try to extend this sort of analysis to pseudo-Riemannian geometries then using the distance function is awkward because it is sometimes real (spacelike separation), sometimes zero (null separation), and sometimes pure imaginary (timelike separation). Synge (and many others) have argued that the best way of taking care of this technical difficulty is to define the "world function" [7]

$$\Omega(x, y) \equiv \frac{1}{2}\, d(x, y)^2 \tag{48}$$

which is always guaranteed to be real (though it can be positive, zero, or negative).

In the present situation we have found it useful to use the notion of characteristic surfaces to encode the causal structure of the system of PDEs given in equation (11) and to define the real quantity $Q(q, p)$ which can be positive, zero, or negative. Now $Q(q, p)$ is not homogeneous linear in p, rather it is homogeneous of order 2N:

$$Q(q, \lambda p) = \lambda^{2N} Q(q, p). \tag{49}$$

If we try to define the analog of the normal Finsler distance function, we would choose

$$d_F(q, p) = [Q(q, p)]^{1/2N}. \tag{50}$$

If $Q(q, p)$ were always positive (which would correspond to an elliptic system of PDEs) this construction can be used to provide a real and positive distance function suitable for defining a Finsler geometry. Unfortunately, we are interested in the hyperbolic case, so $Q(q, p)$ by definition possesses zeros and changes sign. So $d_F(q, p)$ is now generically complex; one typically encounters various branch cuts involving 2N'th roots of -1, $(\exp\{i\pi/(2N)\})$, which invalidates the standard presentation of Finsler geometries. (We emphasize that this is a technical issue, not a fundamental issue, but it does mean one cannot simply copy results from the standard mathematics literature.) Note that even in the case of a unique spacetime metric (completely reducible, so that $Q(q, p) = [Q_0(q, p)]^N$) one still has $d_F(q, p) = \sqrt{Q_0(q, p)} = d_0(q, p)$. In this case the Finsler function degenerates to the usual Lorentzian distance distance function (which is positive real, zero, or pure imaginary). In view of the above, we see that it is $Q(p, q)$ itself that should be thought of as fundamental: it is the natural pseudo-Finslerian generalization of Synge's world function.

There have been several attempts at defining Lorentzian-signature pseudo-Finsler geometries (for example, Asanov [8]). Unfortunately those pseudo-Finsler geometries are typically set up in such a way as to *avoid* the possibility of multiple light cones, which is exactly the situation we are trying to probe in this article. It does not seem to us that the Asanov formulation of pseudo-Finsler geometries has anything to say about the situation at hand. While it is clear that in the general case we will want to invoke some form of pseudo-Finsler geometry, none of the extant formalisms are really suitable [9, 10, 11, 12, 13]. We hope that these brief comments might stimulate some interest in further developing this field.

E. Hidden geometries — Eliminating spurious fields

There are cases in which straightforward application of the previous analysis of characteristics does not work because the generalized Fresnel equation (22) is identically zero. This happens in situations in which one of the fields is spurious, either because of a gauge invariance or possibly because of some non-obvious algebraic (not differential) relation between the fields. However, it should be noted that in some of these cases one can eliminate the spurious field completely, and thereby show that the remaining physical fields are coupled to a "reduced" effective metric.

For example, let us take the general set of equations (11). Consider the situation in which for a particular equation, say equation number 1, and a particular field, say ϕ_1^1, we have ${f^{\mu\nu}}_{11} = 0$, and ${\Gamma^\mu}_{11} = 0$, but $K_{11} \neq 0$. Now choose the notation a, b to denote those indices A, B distinct from 1. Then, the relevant first equation from the system (11) reads:

$$K_{11}\, \phi_1^1 + \partial_\mu \left(f^{\mu\nu}_{1b}\ \partial_\nu \phi_1^b \right) + \frac{1}{2} \left(\Gamma^\mu_{1b}\ \partial_\mu \phi_1^b + \partial_\mu [\Gamma^\mu_{1b}\ \phi_1^b] \right) + K_{1b}\, \phi_1^b = 0. \tag{51}$$

This means that we can *algebraically* solve for the field ϕ_1^1 using

$$\phi_1^1 = -\frac{\partial_\mu\left(f^{\mu\nu}_{1b}\;\partial_\nu\phi_1^b\right) + \frac{1}{2}\left(\Gamma^\mu_{1b}\;\partial_\mu\phi_1^b + \partial_\mu[\Gamma^\mu_{1b}\;\phi_1^b]\right) + K_{1b}\;\phi_1^b}{K_{11}}. \tag{52}$$

This can now be used to eliminate ϕ_1^1 from the system (11) leading to a reduced $[(N-1)\times(N-1)]$ system of equations. To do this we first write the remaining a equations as

$$\partial_\mu\left(f^{\mu\nu}_{ab}\;\partial_\nu\phi_1^b\right) + \partial_\mu\left(f^{\mu\nu}_{a1}\;\partial_\nu\phi_1^1\right) + \frac{1}{2}\left(\Gamma^\mu_{ab}\;\partial_\mu\phi_1^b + \partial_\mu[\Gamma^\mu_{ab}\;\phi_1^b]\right) + \frac{1}{2}\left(\Gamma^\mu_{a1}\;\partial_\mu\phi_1^1 + \partial_\mu[\Gamma^\mu_{a1}\;\phi_1^1]\right) + K_{ab}\;\phi_1^b + K_{a1}\;\phi_1^1 = 0, \tag{53}$$

and then substitute ϕ_1^1.

In general, this reduced set of equations has up to fourth-order derivatives of the linearized fields and so cannot be analyzed along the lines developed in this paper. However, there exist some particular situations in which the reduced set of equations is still second order, though no longer formally self-adjoint. Specifically, within this method of eliminating spurious degrees of freedom, there are three possible particular cases:

- Case a: ${f^{\mu\nu}}_{1b} = 0$ and $\Gamma^\mu_{1b} = 0$,
- Case b: ${f^{\mu\nu}}_{a1} = 0$ and $\Gamma^\mu_{a1} = 0$,
- Case c: ${f^{\mu\nu}}_{1b} = 0$ and ${f^{\mu\nu}}_{a1} = 0$.

In all three of these cases the reduced set of equations can be written in second-order form

$$\partial_\mu\left({\tilde{f}^{\mu\nu}}_{ab}\;\partial_\nu\phi_1^b\right) + \tilde{\Gamma}^\mu_{ab}\;\partial_\mu\phi_1^b + \tilde{K}_{ab}\;\phi_1^b = 0. \tag{54}$$

Specifically, we have:

- Case a:

$$\tilde{f}^{\mu\nu}_{ab} = {f^{\mu\nu}}_{ab} - {f^{\mu\nu}}_{a1}\;\frac{K_{1b}}{K_{11}} \tag{55}$$

$$\tilde{\Gamma}^\mu_{ab} = \Gamma^\mu_{ab} - {f^{\mu\nu}}_{a1}\;\partial_\nu\left(\frac{K_{1b}}{K_{11}}\right) - \Gamma^\mu_{a1}\;\frac{K_{1b}}{K_{11}} \tag{56}$$

$$\tilde{K}_{ab} = K_{ab} - \partial_\mu\left[{f^{\mu\nu}}_{a1}\;\partial_\nu\left(\frac{K_{1b}}{K_{11}}\right)\right] - \Gamma^\mu_{a1}\;\partial_\mu\left(\frac{K_{1b}}{K_{11}}\right) - \frac{1}{2}(\partial_\mu\Gamma^\mu_{a1})\;\frac{K_{1b}}{K_{11}} + \frac{1}{2}(\partial_\mu\Gamma^\mu_{ab}) - K_{a1}\;\frac{K_{1b}}{K_{11}}, \tag{57}$$

- Case b:

$$\tilde{f}^{\mu\nu}_{ab} = {f^{\mu\nu}}_{ab} - {f^{\mu\nu}}_{1b}\;\frac{K_{a1}}{K_{11}}, \tag{58}$$

$$\tilde{\Gamma}^\mu_{ab} = \Gamma^\mu_{ab} + \partial_\nu\left(\frac{K_{a1}}{K_{11}}\right)\;{f^{\nu\mu}}_{1b} - \frac{K_{a1}}{K_{11}}\;\Gamma^\mu_{1b}, \tag{59}$$

$$\tilde{K}_{ab} = K_{ab} - K_{a1}\;\frac{K_{1b}}{K_{11}} - \frac{1}{2}(\partial_\mu\Gamma^\mu_{1b})\frac{K_{a1}}{K_{11}} + \frac{1}{2}(\partial_\mu\Gamma^\mu_{ab}), \tag{60}$$

- Case c:

$$\tilde{f}^{\mu\nu}_{ab} = {f^{\mu\nu}}_{ab} - \frac{1}{2K_{11}}\left(\Gamma^\mu_{a1}\Gamma^\nu_{1b} + \Gamma^\nu_{a1}\Gamma^\mu_{1b}\right), \tag{61}$$

$$\begin{aligned}\tilde{\Gamma}^\mu_{ab} = {} & \Gamma^\mu_{ab} + \frac{1}{2K_{11}}(\partial_\nu\Gamma^\nu_{a1})\Gamma^\mu_{1b} - \frac{1}{2K_{11}}\Gamma^\mu_{a1}(\partial_\nu\Gamma^\nu_{1b}) + \partial_\nu\left[\frac{1}{2K_{11}}\Gamma^\mu_{a1}\Gamma^\nu_{1b} - \frac{1}{2K_{11}}\Gamma^\nu_{a1}\Gamma^\mu_{1b}\right] \\ & -\frac{K_{1b}}{K_{11}}\Gamma^\mu_{a1} - \frac{K_{a1}}{K_{11}}\Gamma^\mu_{1b}\end{aligned} \tag{62}$$

$$\begin{aligned}\tilde{K}_{ab} = {} & K_{ab} - \frac{1}{2}\Gamma^\mu_{ab}\partial_\mu\left(\frac{1}{K_{11}}\partial_\nu\Gamma^\nu_{1b}\right) - \Gamma^\mu_{a1}\partial_\mu\left(\frac{K_{1b}}{K_{11}}\right) - \frac{1}{4K_{11}}(\partial_\mu\Gamma^\mu_{a1})(\partial_\nu\Gamma^\nu_{1b}) \\ & -\frac{K_{1b}}{2K_{11}}(\partial_\mu\Gamma^\mu_{a1}) - \frac{K_{a1}}{2K_{11}}(\partial_\mu\Gamma^\mu_{1b}) - \frac{K_{a1}}{K_{11}}K_{1b}\frac{1}{2}(\partial_\mu\Gamma^\mu_{ab}).\end{aligned} \tag{63}$$

If one now forgets about the initial non-linear system, the two cases that follow the patterns (a) or (b) are actually trivial, in the sense that the field ϕ_1^1 can be seen as an artificial degree of freedom introduced in order to write the equation of motion of a $(N-1)$ fields as an N-component coupled system. Case (c) is a bit more subtle. In fact, the Lagrangian of a irrotational barotropic fluid follows this pattern. We will work out this example explicitly in section VI.

Having completed the reduction, the new system of equations (54) can be analyzed as to its causal and geometric structures along the same lines as before. The new matrices $\tilde{f}^{\mu\nu}{}_{ab}$ will determine the characteristics of the reduced system. Although the reduced system could now fail to be self-adjoint, the previous analysis can nevertheless still be applied — this is because (once spurious fields have been eliminated) the characteristics depend only on the leading symbol of the system of PDEs and are insensitive to lower order terms; equivalently the Fresnel equation determined form the eikonal approximation is insensitive to any possible failure of self-adjointness. Once the behaviour of the fields ϕ_1^b has been determined, the remaining spurious field ϕ_1^1 can (if desired) be obtained from ϕ_1^b by differentiation.

F. Hidden approximate geometries

Another situation of considerable interest occurs when, starting from a system of N coupled PDEs, one can neglect some of the degrees of freedom in some particular limited regime. In this case one can again perform a reduction process, but it is now important to remember that the reduced system is only an approximation to the exact behaviour of the system in the particular regimen.

Let us take units such that time and length have the same dimension (some convenient reference velocity in the system is set to be 1) and the elements of K_{AB} are non-dimensional. Then, the elements of Γ^μ_{AB} must have the dimension of length, and those of the matrices $f^{\mu\nu}{}_{AB}$ the dimension of length squared.

We can now assign to each dimensional coefficient in the set of second-order differential equations (11) a length scale. For instance, thinking of case (c), we might have a situation in which the length scales associated with the terms $f^{\mu\nu}{}_{11}$, Γ^μ_{11}, $f^{\mu\nu}{}_{1b}$, and $f^{\mu\nu}{}_{a1}$ are all much smaller than those of the remaining terms in the system of equations. Let L_s be the largest length scale associated with the terms $f^{\mu\nu}{}_{11}$, Γ^μ_{11}, $f^{\mu\nu}{}_{1b}$, and $f^{\mu\nu}{}_{a1}$, and let L_l be the smallest length scale among all the other terms. Then, provided $L_s \ll L_l$, we can (in dealing with the physics at length scales larger than L_s) safely approximate

$$f^{\mu\nu}{}_{11},\ \Gamma^\mu_{11},\ f^{\mu\nu}{}_{1b},\ f^{\mu\nu}{}_{a1} \simeq 0 \tag{64}$$

In these circumstances, for momenta low in comparison with $1/L_s$, (large wavelengths in comparison with L_s) we will find that the fields ϕ_1^b will behave as if they were coupled to an approximate effective metric determined by

$$\tilde{f}^{\mu\nu}{}_{ab} = f^{\mu\nu}{}_{ab} - \frac{1}{2K_{11}} \left(\Gamma^\mu_{a1}\,\Gamma^\nu_{1b} + \Gamma^\nu_{a1}\,\Gamma^\mu_{1b}\right) \tag{65}$$

This is exactly what happens with the effective Lorentzian metrics arising in acoustic phenomena in dilute gas Bose–Einstein condensates (BECs) [14, 15]. In this case $L_s \sim h/(m\,c_0)$, (where h is Planck's constant, m is the mass of the atoms making up the gas, and c_0 is the velocity of sound in the condensate). This means that the geometric acoustics approximation in BECs is valid for wavelengths larger than $h/(m\,c_0)$ (an "acoustic Compton wavelength") at which one would probe the discrete nature of the condensed gas.

The characteristic surfaces that one can construct based in these low-momentum metrics are only an approximation, and will only make sense when exploring the system with momenta lower that the scale $1/L_s$. By increasing the momentum we will find the true characteristic surfaces of the system. For example, in a BEC the true characteristics are those of a diffusion equation allowing infinite speed propagation. Nevertheless the effective metric found in low energies is a well defined Lorentzian metric with a maximum propagation velocity [16].

G. Fermionic fields

The last possibility we shall comment on in this section is the extreme case in which all the matrices $f^{\mu\nu}{}_{AB}$ are strictly zero. When this happens we cannot directly apply the previously derived analysis in terms of characteristic surfaces. What we should do instead is to either develop a modified discussion based on definition (25), or more prosaically, to first produce a system of second-order PDEs by iteration. Apply the first order operator

$$\hat{D}_{AB}\,\phi_1^B \equiv \frac{1}{2}\left(\Gamma^\mu_{AB}\,\partial_\mu\phi_1^B + \partial_\mu(\Gamma^\mu_{AB}\,\phi_1^B)\right) + K_{AB}\,\phi_1^B, \tag{66}$$

twice. Then

$$\hat{D}_{AC}\, \hat{D}_{CB}\, \phi_1^B = 0. \tag{67}$$

With this procedure we arrive at an equation of the form

$$\partial_\mu \left(\tilde{f}^{\mu\nu}{}_{AB}\, \partial_\nu \phi_1^B \right) + \tilde{\Gamma}^\mu_{AB}\, \partial_\mu \phi_1^B + \tilde{K}_{AB}\, \phi_1^B = 0, \tag{68}$$

with coefficients

$$\tilde{f}^{\mu\nu}{}_{AB} = \frac{1}{2}(\Gamma^\mu{}_{AC}\, \Gamma^\nu{}_{CB} + \Gamma^\nu{}_{AC}\, \Gamma^\mu{}_{CB}), \tag{69}$$

$$\tilde{\Gamma}^\mu{}_{AB} = K_{AC}\, \Gamma^\mu{}_{CB} + K_{BC}\, \Gamma^\mu{}_{CA} + \frac{1}{2}\partial_\nu(\Gamma^\mu{}_{AC}\, \Gamma^\nu{}_{CB} + \Gamma^\nu{}_{AC}\, \Gamma^\mu{}_{CB}) + (\partial_\nu \Gamma^\nu{}_{AC})\, \Gamma^\mu{}_{CB}, \tag{70}$$

$$\tilde{K}_{AB} = \Gamma^\nu{}_{AC}\, \partial_\nu[(\partial_\mu \Gamma^\mu{}_{CB}) + K_{CB}] + [(\partial_\nu \Gamma^\nu{}_{AC}) + K_{AC}]\, [(\partial_\mu \Gamma^\mu{}_{CB}) + K_{CB}]. \tag{71}$$

The system of equations (68) now matches the pattern of our previous discussion.

This situation is characteristic of the existence of fermionic fields. For example in the trivial case of the Dirac equation in flat space, we have $\Gamma^\nu{}_{AC} = \gamma^\nu{}_{AC}$ with AB now denoting spinorial indices. It is easy to see that

$$\tilde{f}^{\mu\nu}{}_{AB} = \frac{1}{2}(\gamma^\mu{}_{AC}\, \gamma^\nu{}_{CB} + \gamma^\nu{}_{AC}\, \gamma^\mu{}_{CB}) = \delta_{AB}\, \eta^{\mu\nu}, \tag{72}$$

so that the different spinorial fields all feel the same Minkowski metric $\eta^{\mu\nu}$.

V. TWO–FIELD SYSTEMS

If we are dealing with a simple two-field system, much of the algebra simplifies. This makes for a useful illustrative example. We start by writing $Q(q,p)$ in terms of a completely symmetric four index tensor, (a quartic)

$$Q(q,p) = Q(q)^{\mu\nu\rho\sigma}\; p_\mu\, p_\mu\, p_\rho\, p_\sigma = \{f^{\mu\nu}{}_{11}(q)\; f^{\rho\sigma}{}_{22}(q) - f^{\mu\nu}{}_{12}(q)\; f^{\rho\sigma}{}_{12}(q)\}\; p_\mu\, p_\mu\, p_\rho\, p_\sigma. \tag{73}$$

In discussing the most general causal structure for the propagation of a two-field system it is this quartic $Q(q)^{\mu\nu\rho\sigma}$ that is the main geometrical object; instead of the metric $f(q)^{\mu\nu} = \sqrt{-g}\; g^{\mu\nu}$ as in the single-field case [1]. We can think of $f^{\mu\nu}{}_{AB}$ as a 2×2 matrix in field space, $f^{\mu\nu}{}_{AB}$, whose components are point-dependent tensors (tensor densities, in fact):

$$f^{\mu\nu}{}_{AB} = \begin{bmatrix} f^{\mu\nu}{}_{11} & f^{\mu\nu}{}_{12} \\ f^{\mu\nu}{}_{12} & f^{\mu\nu}{}_{22} \end{bmatrix}. \tag{74}$$

By redefining the linearized fields we can transform

$$f^{\mu\nu}{}_{AB} \to \bar{f}^{\mu\nu}{}_{AB} = L_A{}^C\, L_B{}^D\, f^{\mu\nu}{}_{CD} \equiv M_{AB}{}^{CD}\, f^{\mu\nu}{}_{AB}. \tag{75}$$

In general, there is no combination of the three independent tensor components of $f^{\mu\nu}{}_{AB}$, even with point-dependent coefficients, that vanishes. That is, typically

$$M_{AB}{}^{11}(x)\; f^{\mu\nu}{}_{11}(x) + M_{AB}{}^{22}(x)\; f^{\mu\nu}{}_{22}(x) + 2M_{AB}{}^{12}(x)\; f^{\mu\nu}{}_{12}(x) \neq 0. \tag{76}$$

In this algebraically most general case, we are forced to deal with pseudo-Finsler geometries. There are other cases in which the three tensor components of $f^{\mu\nu}{}_{AB}$ are not independent. Then, by using field redefinitions we can arrive at five different canonical cases:

I: Two independent components:

Ia: A diagonal component is set to zero —

$$f_{AB} = \begin{bmatrix} f_{11} & f_{12} \\ f_{12} & 0 \end{bmatrix}. \tag{77}$$

In this case

$$Q(q,p) = -\left(f^{\mu\nu}{}_{12}\, p_\mu\, p_\nu\right)^2. \tag{78}$$

This case corresponds to quadratic reducibility for the function $Q(q,p)$ *without diagonalization of the kinetic terms*. There are two induced metrics, both based on $f^{\mu\nu}{}_{12} = \sqrt{-g}\; g^{\mu\nu}$, which identical to each other. However these induced metrics, while they successfully reproduce $Q(q,p)$, and so reproduce the characteristic surfaces, are not enough to reproduce the symbol of the system of PDEs. From the induced metrics we can reconstruct $Q(q,p)$ but not $f^{\mu\nu}{}_{AB}$, so that we cannot arrange to diagonalize the system of PDEs in field space.

Ib: The off-diagonal component is zero —

$$f_{AB} = \begin{bmatrix} f_{11} & 0 \\ 0 & f_{22} \end{bmatrix}. \tag{79}$$

In this case

$$Q(q,p) = \left(f^{\mu\nu}{}_{11}\, p_\mu\, p_\nu\right)\,\left(f^{\mu\nu}{}_{22}\, p_\mu\, p_\nu\right). \tag{80}$$

It is this case that corresponds to a proper bi-metric theory with diagonal kinetic terms and two distinct metrics.

II: There is only one independent component:

IIa: Only one diagonal component is different from zero —

$$f_{AB} = \begin{bmatrix} f_{11} & 0 \\ 0 & 0 \end{bmatrix}. \tag{81}$$

In this case $Q(q,p) \equiv 0$; this is a sign that one of the fields is unphysical, either because it is superfluous or because it is actually a gauge degree of freedom.

IIb: The two diagonal components are different from zero — and equal

$$f_{AB} = \begin{bmatrix} f_{11} & 0 \\ 0 & f_{11} \end{bmatrix}. \tag{82}$$

This is a true single-metric theory with diagonalizable kinetic energy terms and

$$Q(q,p) = -\left(f^{\mu\nu}{}_{11}\, p_\mu\, p_\nu\right)^2. \tag{83}$$

This is the case compatible with strict adherence to the Einstein Equivalence Principle.

III: All the tensor components of f_{AB} vanish:
This probably means you are dealing with a Fermi field and somehow did not notice. See section (IV G) above.

In the three cases I_a, I_b and II_b the quartic (73) factorizes into the product of two quadratics. The case I_b corresponds to a proper bi-metric theory in which there are two fields each reacting to a different metric. In the cases I_a and II_b the two fields feel the same metric and so we would recover the usual behaviour of classical fields (such as Maxwell or Dirac fields) in general relativity. Of course in all these cases it is still necessary to analyze if the relevant metrics are Lorentzian. (And to verify that they satisfy suitable causality constraints.) The case II_a is more tricky: The quartic (73) is degenerate and so we cannot strictly define any kind of characteristic surface. However, as we have previously argued, in some circumstances it is still possible to reduce the number of fields and so define characteristic surfaces based in the existence of a hidden metric. We shall now demonstrate this and related phenomenon explicitly with a pair of simple examples.

VI. EXAMPLES

A. The barotropic irrotational inviscid fluid

We can write the Lagrangian for a barotropic inviscid irrotational fluid as [17, 18]

$$\mathcal{L} = \frac{1}{2}\rho(\boldsymbol{\nabla}\theta)^2 + \rho\,\partial_t\theta + \int_0^\rho d\rho'\; h(\rho'). \tag{84}$$

Here, the two fields $\phi^A \equiv (\rho,\, \theta)$ are the fluid density ρ and the velocity potential θ (the velocity field $\mathbf{v}$ can be written as $\mathbf{v} = \boldsymbol{\nabla}\theta$ because the irrotational nature of the fluid considered). The function

$$h(\rho) = h[p(\rho)] = \int_0^p \frac{dp'}{\rho(p')} \tag{85}$$

is the enthalpy of the fluid. Now, by linearizing we obtain the relevant second-order Lagrangian for the perturbations

$$\mathcal{L}^{(2)} = \frac{1}{2}\rho_0(\boldsymbol{\nabla}\theta_1)^2 + \rho_1 \boldsymbol{\nabla}\theta_1 \cdot \boldsymbol{\nabla}\theta_0 + \rho_1 \partial_t \theta_1 + \frac{1}{2\rho_0} \frac{dp}{d\rho}\bigg|_0 \rho_1^2. \tag{86}$$

(Remember that the linear terms cancel by using the background equation of motion). From this we get the two linearized equations of motion

$$-\left(\partial_t\theta_1 + \mathbf{v_0}\cdot\boldsymbol{\nabla}\theta_1 + \frac{c_0^2}{\rho_0}\,\rho_1\right) = 0, \tag{87}$$

$$\partial_t\rho_1 + \boldsymbol{\nabla}(\rho_0\boldsymbol{\nabla}\theta_1) + \boldsymbol{\nabla}(\rho_1\mathbf{v_0}) = 0, \tag{88}$$

in which we have used the definitions $\mathbf{v_0} \equiv \boldsymbol{\nabla}\theta_0$ and $c_0^2 \equiv (dp/d\rho)|_0$. (Of course, these must and do agree with the equations obtained by the more traditional means of first finding the full equations of motion and then linearizing.) As we can see, equation (87) has only the zero-order term for ρ_1 and in the discussion of hidden geometries the condition *(c)* above is fulfilled. In our previous notation (note that $\boldsymbol{\rho}$ and $\boldsymbol{\theta}$, which we boldface for clarity, are now field indices not spacetime indices)

$$f^{\mu\nu}{}_{\boldsymbol{\rho\rho}} = f^{\mu\nu}{}_{\boldsymbol{\rho\theta}} = f^{\mu\nu}{}_{\boldsymbol{\theta\rho}} = 0, \tag{89}$$

$$f^{\mu\nu}{}_{\boldsymbol{\theta\theta}} = \begin{bmatrix} 0 & \vdots & 0 \\ \cdots & \cdot & \cdots \\ 0 & \vdots & \rho_0\,\delta^{ij} \end{bmatrix}, \tag{90}$$

and

$$\Gamma^{\mu}{}_{\boldsymbol{\rho\rho}} = \Gamma^{\mu}{}_{\boldsymbol{\theta\theta}} = 0, \tag{91}$$

$$\Gamma^{\mu}{}_{\boldsymbol{\theta\rho}} = \{1, v_0^i\}, \tag{92}$$

$$\Gamma^{\nu}{}_{\boldsymbol{\rho\theta}} = \{-1, -v_0^i\}, \tag{93}$$

while

$$K_{\boldsymbol{\theta\theta}} = K_{\boldsymbol{\rho\theta}} = K_{\boldsymbol{\theta\rho}} = 0, \tag{94}$$

$$K_{\boldsymbol{\rho\rho}} = -\frac{c_0^2}{\rho_0}. \tag{95}$$

Finally, by looking at case (c) as given in equation (63) we can calculate all the terms in equation (54). We have $\tilde{K} = 0$, $\tilde{\Gamma} = 0$ and a *single* reduced metric of the form

$$\tilde{f}^{\mu\nu} = \frac{\rho_0}{c_0^2} \begin{bmatrix} -1 & \vdots & -v_0^i \\ \cdots & \cdot & \cdots\cdots\cdots\cdots \\ -v_0^i & \vdots & c_0^2\delta^{ij} - v_0^i\, v_0^j \end{bmatrix}, \tag{96}$$

which corresponds to the usual acoustic metric[19, 20]. Once we have the solutions for θ_1, the solutions for ρ_1 can be found by substituting in (87).

On the other hand, if we tried to eliminate θ_1 to find an equation from which to derive the same set of solutions for ρ_1, we would find a very involved equation. Formally one can write

$$\rho_1 = \Gamma^{\mu}{}_{\boldsymbol{\rho\theta}}\, \partial_\mu\theta_1 = \mathcal{D}\,\theta_1, \tag{97}$$

which implies

$$\theta_1 = \mathcal{D}^{-1}\rho_1. \tag{98}$$

Then we have

$$\partial_\mu (f^{\mu\nu} \partial_\nu \theta_1) = 0, \tag{99}$$

and

$$\mathcal{D} \partial_\mu (f^{\mu\nu} \partial_\nu \mathcal{D}^{-1} \rho_1) = 0. \tag{100}$$

While at low momentum the differential equation satisfied by ρ_1 is rather complicated, we can see that in the large momentum limit (the eikonal approximation) the propagation of density waves sees the same causal structure as the θ_1-waves. That is, the causal structure for the propagation of both fields is described by the same metric determined by $f^{\mu\nu} = \sqrt{-g}\, g^{\mu\nu}$.

Of course with hindsight the reason this elimination procedure works is that it is possible to go all the way back to the Lagrangian in equation (84) and eliminate the density from the Lagrangian before linearizing. Nevertheless, this is a nice simple example of how to eliminate spurious fields from the linearized system.

B. The BEC analog model

Let us consider now a Bose–Einstein condensate. It can be described by the Gross–Pitaevskii equation,

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \psi(t, \mathbf{x}) = \left(-\frac{\hbar^2}{2m} \nabla^2 + V_{\rm ext}(\mathbf{x}) + \lambda \ |\psi(t, \mathbf{x})|^2 \right) \psi(t, \mathbf{x}), \tag{101}$$

where $\psi(t, \mathbf{x})$ represents the mean field wave function. By using the Madelung representation [21]

$$\psi(t, \mathbf{x}) = \sqrt{\rho(t, \mathbf{x})}\ \exp[-\mathrm{i}m\theta(t, \mathbf{x})/\hbar], \tag{102}$$

one can separate the (complex) Gross-Pitaevskii equation into the two (real) equations:

$$\partial_t \theta + \frac{1}{2} (\boldsymbol{\nabla} \theta)^2 + \frac{V_{\rm ext}}{m} + \frac{\lambda}{m} \rho - \frac{\hbar^2}{2m^2} \ \left(\frac{\Delta \sqrt{\rho}}{\sqrt{\rho}} \right) = 0. \tag{103}$$

$$\partial_t \rho + \boldsymbol{\nabla} \cdot (\rho\ \boldsymbol{\nabla} \theta) = 0. \tag{104}$$

Here, the quantity

$$V_{\rm Q}(\rho) \equiv -\frac{\hbar^2}{2m} \left(\frac{\Delta \sqrt{\rho}}{\sqrt{\rho}} \right) \tag{105}$$

is what is commonly called the “quantum potential” [22, 23]. Expanding around a background solution of these equations one finds the following two equations for the perturbed quantities ρ_1 and θ_1

$$-\partial_t \theta_1 - \mathbf{v_0} \cdot \boldsymbol{\nabla} \theta_1 - \frac{c_0^2}{\rho_0} \rho_1 + \frac{\hbar^2}{2m^2}\ D_2 \rho_1 = 0. \tag{106}$$

$$\partial_t \rho_1 + \boldsymbol{\nabla}(\rho_0 \boldsymbol{\nabla} \phi_1) + \boldsymbol{\nabla}(\rho_1 \mathbf{v_0}) = 0 \tag{107}$$

We have used the definitions $\mathbf{v_0} \equiv \boldsymbol{\nabla} \phi_0$ and $c_0^2 \equiv (\lambda/m)\rho_0$. Additionally, D_2 represents a relatively messy second-order differential operator obtained from linearizing the quantum potential. Explicitly:

$$D_2\, \rho_1 \equiv -\frac{1}{2} \rho_0^{-3/2}\ [\Delta(\rho_0^{+1/2})]\ \rho_1 + \frac{1}{2} \rho_0^{-1/2}\ \Delta(\rho_0^{-1/2} \rho_1). \tag{108}$$

These linearized equations are formally equivalent to the ones obtained for the barotropic irrotational fluid except for the presence of the quantum potential. This difference amounts to the existence of the following additional non-zero coefficients:

$$f^{\mu\nu}{}_{\boldsymbol{\rho}\boldsymbol{\rho}} = \frac{\hbar^2}{2m^2} \left[\begin{array}{ccc} 0 & \vdots & 0 \\ \cdots & \cdot & \cdots \\ 0 & \vdots & \frac{1}{2}\rho_0^{-1} \delta^{ij} \end{array} \right], \tag{109}$$

$$\Gamma^\mu{}_{\boldsymbol{\rho}\boldsymbol{\rho}} = \frac{\hbar^2}{2m^2} \left\{ 0,\ \rho_0^{-1/2} \partial^i \rho_0^{-1/2} \right\}, \tag{110}$$

$$K_{\boldsymbol{\rho}\boldsymbol{\rho}} = \frac{\hbar^2}{2m^2} \left[-\rho_0^{-3/2}\ \Delta(\rho_0^{+1/2}) + \rho_0^{-3/2} \boldsymbol{\nabla} \rho_0^{+1/2} \cdot \boldsymbol{\nabla} \rho_0^{+1/2} \right]. \tag{111}$$

Now, following the discussion in the "Hidden approximate geometries" subsection IV F, one can make the K coefficient in equation (106) non-dimensional by multiplying the whole equation by ρ_0/c_0^2. Then, the important point from our example is that it is easy to see that the new terms that occur in the BEC system (with respect to the previous fluid system) all come naturally multiplied by the the length scale $L_s = \hbar/(m\, c_0)$. This corresponds to an "acoustic Compton wavelength" defined by the propagation speed of phonons in the condensate. If one is probing the BEC system at large length scales (low momenta) in comparison with this Compton wavelength, one can neglect the new terms (109), (110), and (111). One can then perform the same trick as with the barotropic irrotational fluid to obtain an effective metric for ρ_1. If we don't neglect the new terms coming from the quantum potential we can still arrive at an integral-differential equation for ρ_1. This equation will encode the Bogoliubov dispersion relation [14, 15], showing us that the symbol of the Gross–Pitaevskii equation is that of a (complex) diffusion equation. As such it allows the propagation of signals at arbitrarily large speeds [16].

VII. SUMMARY AND DISCUSSION

Taking an arbitrary system of hyperbolic second-order PDEs (either generic or arising from a first-order Lagrangian), the behaviour of the perturbations of the fundamental fields around any background configuration can be given a geometrical interpretation. For a single field there is always a nice and clean geometrical interpretation in terms of the d'Alembertian operator in an effective (typically curved) Lorentzian geometry [1]. For several coupled fields, as discussed in this article, the situation is seen to be more complex: In simple cases all the fields see the same effective Lorentzian metric; with a little less luck the fields see different effective Lorentzian metrics, up to one effective metric per field; with extremely bad luck you will need to adopt some Finsler-like extension of the notion of Lorentzian geometry. Taking into account this formal analysis, in order to obtain an analog model of general relativity from a general system of PDEs it will be necessary to first require that the system fulfill some sort of "Einstein Equivalence Principle" (at least approximately). That is, at low energies we would desire an (approximate) factorization:

$$f^{\mu\nu}{}_{AB} \approx h_{AB}\; f^{\mu\nu} = h_{AB}\; \sqrt{-g}\; g^{\mu\nu}. \tag{112}$$

Then, and only then, will all the low-energy fields see (approximately) the same effective Lorentzian metric, as is experimentally implied by the Eötvös experiment. At this level, we have not found it possible to obtain the "Einstein Equivalence Principle" from more fundamental principles.

In searching for a geometrical interpretation coming from a set of PDEs we have analyzed the physical content of the matrix f_{AB}. Sometimes, however, this matrix is singular and does not provide a geometrical interpretation for the behaviour of all of the fields. In these situations one can still find such geometrical interpretation by extracting additional information from the $\Gamma^\mu{}_{AB}$ and K_{AB} coefficients. An extreme example is provided by a Fermi system. In this case, all the $f^{\mu\nu}_{AB}$ are zero but we can find a geometrical structure coming from the $\Gamma^\mu{}_{AB}$. In other situations, it can happen that the system has some "spurious" degrees of freedom making difficult the search for a geometrical interpretation in terms of the $f^{\mu\nu}{}_{AB}$ alone. In many situations we can eliminate these spurious degrees of freedom and arrive to a reduced set of PDEs with a well defined geometrical interpretation. Additionally, there can be situations in which the elimination of spurious degrees of freedom can be justified in an approximate way. By this we mean that even though, strictly speaking, the f_{AB} are non-singular and one should find a geometrical interpretation directly for the complete set of them, nevertheless for some particular regime (energy scale) the f_{AB} can be considered to be approximately singular and an elimination of spurious degrees of freedom appropriate. The approximate geometric structure found by this procedure, though extremely useful, can be completely different from the exact geometric structure based on the exact characteristics. We have illustrated this point with the BEC analog model. The exact underlying geometric structure provided by the non-relativistic Gross–Pitaevskii equation is "parabolic" while the approximate effective geometric structure is Lorentzian or "hyperbolic".

In this paper we have discussed the different possibilities one can encounter in extracting a geometric structure from a system of PDEs. The particular systems that fulfill the Einstein Equivalence Principle can most easily be seen as analog models *of* General Relativity. Without additional structure they should not be thought of as models *for* general relativity, as the dynamics of the effective geometry can differ greatly from the proper general relativistic dynamics (the Einstein equations). On the other hand, in a previous work [1], and the in the context of a single field system, we showed that one-loop quantum effects could also provide, in some circumstances, a dynamics somewhat resembling that of general relativity. We leave as a future project to analyze in detail this quantum mechanism in the context of a multi-field system.

In summary: The occurrence of effective metrics and effective geometries in low-energy linearized approximations to wide classes of dynamical systems is striking — the near ubiquitous occurrence of this effect is particularly useful when considering analog models *of* general relativity, can even be interpreted as being strongly suggestive that some form of "induced gravity" may underly true physical gravity. To turn "induced gravity" into a serious contender

several additional conditions must first be met: On the one hand we will need to be dealing with multiple background fields (at least six) in order to have an algebraically general spacetime metric [1], on the other hand all the linearized background fields must see the "same" Lorentzian metric (either exactly the same metric if you believe strongly in the Einstein Equivalence Principle, or at the very least approximately the same metric in order to be compatible with experimental constraints from Eötvös-type experiments). Only once these two basic conditions have been met will it be sensible to develop a Sakharov-like "induced gravity" [24]. One-loop effects will then generate a term in the effective action that is proportional to the Einstein–Hilbert action [1]. The result would then be an "effective" theory of gravity in the sense of Donoghue [25], but would still suffer from the potential defects common to all embedding models of general relativity [1, 26, 27].

In this regard the results we report in this paper are mixed: The good news is that the occurrence of Lorentzian metrics seems quite common, the bad news is that *multiple* Lorentzian metrics (and worse) seems generic. We have been able to say quite a bit about the general way in which one might probe the causal structure of these theories, and while we have been able to put much of the discussion of multi-refringence in general framework. We are continuing to work on seeking field theories that are more general than simple "minimum substitution" theories, but are still restricted enough to produce a well-controlled metric structure.

Acknowledgements

The research of Matt Visser was supported by the US Department of Energy. Stefano Liberati was supported by the US National Science Foundation. Carlos Barceló was initially supported by the Spanish Ministry of Science and Technology (MCYT), and is currently supported by a European Community Marie Curie grant. The authors wish to thank Sebastiano Sonego for his comments and observations.

[1] C. Barceló, S. Liberati, and M. Visser, "Analog gravity from field theory normal modes?", Classical and Quantum Gravity **18** (2001) 3595-3610 [gr-qc/0104001];
"Einstein gravity as an emergent phenomenon?", gr-qc/0106002; International Journal of Modern Physics (in press).

[2] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol II*, Wiley, John and Sons, (1990).

[3] *Encyclopedic Dictionary of Mathematics*, K. Ito, K. Ito and N.S. Ugakkai (Editors), 2nd ed, MIT Press (1987).

[4] M. Born and E. Wolf, *Principles of optics*, (Pergamon, Oxford, 1980).

[5] L. D. Landau, E. M. Lifshitz and L. P. Pitaevskii, *Electrodynamics of Continuous Media*, 2nd edition (Oxford, Pergamon Press, 1984).

[6] J. Hadamard, *Leçons sur la propagation des ondes et les équations de l'hydrodynamique*, (Hermann, Paris, 1903).

[7] J. L. Synge, *Relativity: the general theory*, (North Holland, Amsterdam, 1960).

[8] G.S. Asanov, *Finsler geometry, relativity, and gauge theories*, (Reidel, Dordrecht, Netherlands, 1985).

[9] A. Bejancu, *Finsler geometry and applications*, (Ellis Horwood, Chichester, England, 1990).

[10] D. Bao, S.S. Chern, and Z. Shen, *An introduction to Riemann–Finsler geometry*, (Springer, New York, 2000).

[11] H. Rund, *The differential geometry of Finsler spaces*, (Springer, Berlin, 1959).

[12] P.L. Antonelli and B.C. Lackey (editors), *The theory of Finslerian Laplacians and applications*, (Kluwer, Dordrecht, 1998).

[13] D. Bao, S.S. Chern, and Z. Shen (editors), *Finsler Geometry*, Proceedings of the Joint Summer Research Conference on Finsler Geometry, July 1995, Seattle, Washington, (American Mathematical Society, Providence, Rhode Island, 1996).

[14] C. Barceló, S. Liberati, and M. Visser, "Analog gravity from Bose–Einstein condensates", Class. Quant. Grav. **18** (2001) 1137 [gr-qc/0011026];
"Towards the observation of Hawking radiation in Bose–Einstein condensates", arXiv:gr-qc/0110036.

[15] L. J. Garay, J. R. Anglin, J. I. Cirac and P. Zoller, "Black holes in Bose–Einstein condensates", Phys. Rev. Lett. **85**, 4643 (2000) [gr-qc/0002015];
"Sonic black holes in dilute Bose–Einstein condensates", Phys. Rev. **A63**, 023611 (2001) [gr-qc/0005131].

[16] M. Visser, C. Barceló, and S. Liberati, "Acoustics in Bose-Einstein condensates as an example of broken Lorentz symmetry", [hep-th/0109033].

[17] R. Jackiw and A. P. Polychronakos, "Supersymmetric fluid mechanics", Phys. Rev. D **62**, 085019 (2000).

[18] L. O'Raifeartaigh and V. V. Sreedhar, "The maximal kinematical invariance group of fluid dynamics and explosion-implosion duality", [hep-th/0007199].

[19] W.G. Unruh, "Experimental black hole evaporation?", Phys. Rev. Lett. **46**, 1351 (1981);
"Dumb holes and the effects of high frequencies on black hole evaporation", Phys. Rev. D **51**, 2827 (1995) [gr-qc/9409008].

[20] M. Visser, "Acoustic black holes: Horizons, ergospheres, and Hawking radiation", Class. Quantum Grav. **15**, 1767 (1998) [gr-qc/9712010];
"Acoustic propagation in fluids: An Unexpected example of Lorentzian geometry", gr-qc/9311028;
"Acoustic black holes", gr-qc/9901047.

[21] E. Madelung, "Quantentheorie in hydrodynamischer Form", Zeitschrift für Physik **38**, 322 (1926).

[22] D. Bohm, "A suggested interpretation of the quantum theory in terms of "hidden" variables: I and II", Phys. Rev. **85**, 166 (1952); **85**, 180 (1952).

[23] P. R. Holland, *The quantum theory of motion: an account of the de Broglie-Bohm causal interpretation of quantum mechanics*, (Cambridge, England, 1993)

[24] A. Sakharov, "Vacuum quantum fluctuations in curved space and the theory of gravitation", Soviet Physics Doklady, **12**, 1040 (1968). [Dokl. Akad. Nauk Ser. Fiz. **177** (1968) 1040].

[25] J. F. Donoghue, "Quantum general relativity is an effective field theory", PRINT-96-198 (MASS.U.,AMHERST) *Talk presented at the 10th International Conference on Problems of Quantum Field Theory, Alushta, Ukraine, 13-17 May 1996*;
"Gravity and Effective Field Theory: A Talk for Phenomenologists", hep-ph/9512287;
"Introduction to the Effective Field Theory Description of Gravity", gr-qc/9512024;
"The Ideas of gravitational effective field theory", hep-th/9409143;
"General relativity as an effective field theory: The leading quantum corrections", Phys. Rev. D **50** (1994) 3874 [gr-qc/9405057].

[26] T. Regge and C. Teitelboim, "General Relativity a la string: a progress report", in *Proceedings of the Marcel Grossman meeting*, Trieste (1975).

[27] S. Deser, F.A.E. Pirani and D.C. Robinson, "New embedding model of general relativity", Phys. Rev. **D14**, 3301 1976).